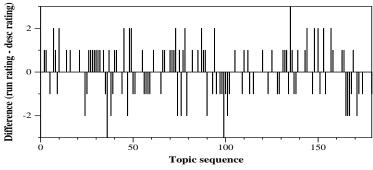
Summary Statistics			
Run ID	bartpodcasts		
Task	Summarization		
Run type	automatic		
Summarization model	Supervised neural or deep learning model; Transfer learn-		
	ing or other pre-trained model		
Summarization method	Abstractive		
Number of summaries	179		

Overall measures	
Summaries rated "excellent"	25 (14.0%)
Summaries rated "good"	50 (27.9%)
Summaries rated "fair"	50 (27.9%) 67 (37.4%)
Summaries rated "bad"	37 (20.7%)

ROUGE score compared to episode description		
ROUGE_L P score	0.2078 (95%-conf.int. 0.19668 - 0.21971)	
ROUGE_L R score	0.2101 (95%-conf.int. 0.20063 - 0.22000)	
ROUGE_L F score	0.1663 (95%-conf.int. 0.15924 - 0.17371)	

Qualitative judgments		
Q1: Does the summary include names of the main people	51.4%	
(hosts, guests, characters) involved or mentioned in the		
podcast?		
Q2: Does the summary give any additional information	34.6%	
about the people mentioned (such as their job titles, bi-		
ographies, personal background, etc)?		
Q3: Does the summary include the main topic(s) of the	78.2%	
podcast?		
Q4: Does the summary tell you anything about the format	49.2%	
of the podcast; e.g. whether it's an interview, whether it's		
a chat between friends, a monologue, etc?		
Q5: Does the summary give you mode context on the title		
of the podcast?		
Q6: Does the summary contain redundant information?		
Q7: Is the summary written in good English?		
Q8: Are the start and end of the summary good sentence		
and paragraph start and end points?		



Per-topic difference in rating from episode description